DOCUMENT RESUME

ED 068 968 24 CS 200 218

AUTHOR Williams, Robert I.; Nattinger, James R.
TITLE Determination of Criteria to Measure English Language Responsiveness. Final Report.
INSTITUTION Oregon State Univ., Portland.
SPONS AGENCY Office of Education (DHEW), Washington, D.C. Regional Research Program.
BUREAU NO BR-1-J-011
PUB DATE Aug 72
CONTRACT OEC-X-71-0023(057)
NOTE 53p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *College Students; Deep Structure; Dialects; Discourse Analysis; Grammar; Language Ability; *Language Research; *Language Skills; *Linguistics; Semantics; Sentences; Sociolinguistics; Syntax; Test Reliability; Tests; Test Validity; *Written Language

ABSTRACT

The purpose of this study was to locate those measurable criteria of responsiveness to language which have not been utilized to any significant extent in standardized tests. The project involved creating a testing instrument, reviewing of the instrument by a panel of experts, and applying the resultant test to a small population of college students and to a criterion group of acknowledged sensitive users of English. By means of pre- and post-test, comparison was made between the performances of students and members of the criterion group. Reliability and validity studies were made of the results. Though the test was necessarily crude, it was found that certain areas of linguistic response warranted serious consideration in the development of future language competency tests: sociolinguistic and semantic context, deep structure of word-groups and sentences, primary factors of extended discourse, and, possibly, organization of extended discourse. (Author)

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED 066 733

Final Report

Project No. 1-J-011
Contract No. OEC-X-71-0023(057)

DETERMINATION OF CRITERIA TO MEASURE ENGLISH LANGUAGE RESPONSIVENESS

Robert I. Williams
James R. Nattinger
English Department

Portland State University

P.O. Box 751
Portland, Oregon
97207

August, 1972

The research reported herein was performed pursuant to a contract with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgement in the conduct of the project. Points of view or opinions do not, therefore, necessarily represent official Office of Education position or policy.

U.S. Department of
Health, Education, and Welfare

Office of Education

Regional Research Program

FILMED FROM BEST AVAILABLE COPY

ABSTRACT

The purpose of the study was to locate measurable criteria of responsiveness to language which have not been utilized to any significant extent in standardized tests. The project involved creating a testing instrument, reviewing of the instrument by a panel of experts, and applying the resultant test to a small population of college students and to a criterion group of acknowledged sensitive users of English. Comparison was made between the performances of students and members of the criterion group by means of a pre-and-post test. Reliability and validity studies were made of the results. Though the test was necessarily crude, it was found that certain areas of linguistic response warranted serious consideration in the development of future language competency tests: sociolinguistic and semantic context, deep structure of word-groups and sentences, primary factors of extended discourse, and, possibly, organization of extended discourse.

INTRODUCTION

Under a contract awarded by Region X of the U.S. Office of Education, Department of Health, Education, and Welfare members of the English Department of Portland State University, Portland, Oregon, conducted a pilot research project to isolate measurable criteria of responsiveness to language irrespective of dialect. Research was confined to written English in a variety of forms. The project was conceived because the researchers felt that most standardized language competency tests used in schools and colleges were superficial, that they measured the ability of the student to know and use certain expected forms of standard English, but seldom reached other linguistic abilities. Responsiveness--the capacity to recognize, if not intellectualize, nuances of word meaning, of structural form, and of rhetorical organization--was the term we used to represent those linguistic and sociolinguistic capacities left largely unmeasured by existing tests. Though the term was not entirely satisfactory, it did indicate the complexity and multiplicity of individual responses to various dialects in a speech community, e.g., the person who does not respond fully to the dialects of the university or business world but may react with great sensitivity to the dialect of his own group. To what extent this multiplicity of response was due to environmental factors we could not know, for while studies in the language of the disadvantaged child emphasize his inability to handle complex linguistic forms (Bernstein 1961, Deutsch 1963), few have probed deeply enough to determine whether this inability is merely symptomatic of a particular sociocultural environment (Labov 1969a, Wolfram 1969). Regardless of the ultimate causes of inadequacies with some particular linguistic forms, specialists generally recognize that the formal requirements of "good English" do not really reflect native capacity to understand and use another dialect of English (Shuy 1964, Labov 1969b, Baratz and Shuy 1969, Burling 1970). At the same time, no one has adequately determined, much less measured, the constituents of this native capacity.

Measurement of language competence is difficult; it is even more difficult when it attempts to quantify the elements of an intuitive ability which amounts to sensitivity to a number of complex linguistic factors. There has been some research on developing instruments to register linguistic aptitude (Carroll 1954, 1961, Hunt 1968, Tikofsky 1968), but evaluations of objective tests which attempt to measure

"aptitude," "competency," "communicational skill," etc. show that the tests are all but inadequate. In a survey of evaluations of sixteen standard tests of linguistic ability in junior high and high school the researcher concludes, "the evidence of this study points overwhelmingly to the fact that there is no perfect objective test of English, nor does any currently published test come close to the goal of measuring success in English" (Wood 1968). Not only are tests inadequate, criteria by which evaluations are made are evidently highly variable (Smith 1969). Such data not only point to the difficulty of objective testing with language, they suggest that existing tests may not be registering significant features of language response. From our own survey of testing materials we felt that they tended to emphasize superficial competencies at the expense of linguistic ability. This tendency, along with the difficulty of language measurement in general, may account for the but partial success of standardized tests in English.

What standardized tests lacked, we observed, were items which registered a "deep" response to language in the same sense that the linguist distinguishes between "deep" and "surface" structure. This "deep" response might include sensitivity to underlying grammatical structures, influence of speaker and situation on choice of words and structures, nuances of meaning, and subtle relationships operating beneath the surface of extended discourse. Though subjective in nature, responses to such factors needed to be isolated and examined to see whether or not they were measurable. This project was a beginning step in that process, with the specific aim of subjecting a number of likely responses--chiefly the ability to respond to underlying grammatical relationships--to a testing procedure which would indicate the probability of their being measured, given further work and more developed instruments than we would be able to use.

METHODS AND PROCEDURES

The project's principal investigator was Robert I. Williams, a teacher of writing and literature specializing in Renaissance comedy and stylistic analysis; its co-investigator was James R. Nattinger, a linguist specializing in bilingualism and sociolinguistics; its consultant was Morris Weitman, a psychologist specializing in tests and testing procedures. The project took from June of 1971 to June of 1972 not counting time needed for writing of this report. It consisted of selection of criteria to be tested, creation of a testing instrument, application of the instrument to a criterion group and to a subject population of junior and senior college students enrolled in 6 sections of Writing 323 (Junior Composition) in the academic year 1971-1972, and, finally, compilation and interpretation of results. While the results of the project were more indicative than final--we could not develop a fully reliable test, nor could we apply what we did develop to a large population--we succeeded in isolating 6 significant criteria of responsiveness which have some probability of yielding measurable results in future, more extensive projects. Further, the techniques we used may be of value to anyone continuing this line of research.

The first step in the project was selection of potential criteria for analysis and testing. We decided that we would have to be guided by two general principles: that responsiveness, whatever it may be, would involve a somewhat heightened sensitivity to the deep factors operating in semantic meaning, structure of word groups and sentences, and in extended discourse; and second, that our choice of components would be limited, necessarily, by testing. Whether or not the final resultant group of measurable criteria were the most significant as indicators of responsiveness would need to await a considerable advance of general knowledge of linguistics, psycholinguistics, sociolinguistics, and of what specific capacities are likely to make one person a good poet, another no poet at all, or one person facile with words, another not. We composed rough test questions, then, on a number of specific aspects of the three general areas, semantic meaning, structure of word group and sentence, and extended discourse, emphasizing choices that would reveal a response to less than obvious factors operating in a particular case. James Nattinger devised the items on

structure of word groups and sentences, Robert Williams devised those on semantic meaning and extended discourse; each read and criticised each other's items, rejecting those that were unclear or trivial.

Beginning with a stock of rough items totalling 135, we turned to the three problems they presented: were they significant, were they measurable, and were they susceptible to analysis? Whether or not an item was significant, both in itself as a meaningful choice and in respect to the linguistic feature it was meant to test, was determined by a group of five experts in the fields of English writing and linguistics. We chose people with experience in more than one phase of language study and teaching, who were acquainted with the concept of deep structure, and who were, otherwise, of varied backgrounds and interests:

Gregory F. Goekjian, Asst. Prof. English, Ph.D. Pittsburgh, 1970; specialist in literary criticism, poetry.

James R. Nattinger, Asst. Prof. English, Ph.D. Michigan, 1970; linguist, specializing in bilingualism and sociolinguistics.

Marjorie B. Nelson, Assoc. Prof. English, M.A. Oregon, 1958; specialist in Milton, English teacher education; head of reading, English section, E.T.S., Princeton, N.J.

Shelley C. Reece, Asst. Prof. English, Ph.D. Nebraska, 1966; head of Composition Program, PSU, specialist in composition, teaching of writing, rhetoric.

Robert I. Williams, Assoc. Prof. English, Ph.D. California (Berkeley), 1967; specialist in Renaissance comedy, stylistic analysis.

Baxter D. Wilson, Assoc. Prof. English, Ph.D. Virginia, 1952; linguist, specializing in phonetics, Old English.

Note: Nattinger and Williams read each other's test items and factored them independently, jointly making up the fifth member of the group.

Agreement of 4 out of 5 of the group was required to establish that a potential test item was intelligible and significant. Along with agreement on significance, we required the group of experts to have a similar community of opinion on each item's linguistic components, that is, we wanted to assure that each item could be analyzed, either as a choice that would measure a specific component (e.g., semantic meaning) or that would measure two or more components interacting together (e.g., semantic meaning, word form, structural feature). This procedure was necessary, we felt, because a given language choice often has more than one linguistic factor operating in it, and if it were possible to correlate responses we might find that a subject tended to be strong in choices involving, say, structure of word groups and sentences, and weak in those involving semantic meaning.

Agreement among the experts on the linguistic components of each test item was achieved by use of sheets which had a number of "factors" which could enter into the choice a test item called for. Though these factors were somewhat crude, they were sufficiently detailed to indicate whether or not a given test item's response could be analyzed and what linguistic components, specifically, the choice involved. (For a sample factoring sheet, see Appendix A). After discussing these components with the group, and providing them with a glossary so that everyone was clear as to what was meant by the headings and sub-headings on the sheet, Robert Williams and James Mattinger joined the group as experts analyzing each other's test items. Each potential item, then, of the original collection of 135 was subjected to independent analysis by the group. Each of the five experts was asked to distribute 10 points for each item among the various components he felt were operant in the choice the item called for. Thus the test item (1-C), which asked the subject to fill in a blank by choosing between "a real persuasive" and "a really persuasive," had the following distribution of factors according to one expert:

2 Sociolinguistic Context
3 Semantic Context
2 Word Groups
1 Position, Positional Emphasis
2 Underlying Grammatical Relationships

The numbers--parts of 10 distributed--indicate this expert's opinion of the linguistic factors operating in this particular language choice and their relative weights. Such a wide distribution of factors was, however, unusual; typically, the experts saw 2 or 3 factors operating in choices, e.g., an expert saw in a choice having to do with word-compounding (18-A) linguistic factors of the relative weights of 4 for Semantic Context, 6 for Underlying Grammatical Relationships. Where there was not substantial agreement among the experts on the components involved the item was excluded from the test. 80% or more agreement was required. By this process we arrived at a collection of 28 test items, many with two or more parts, totaling 74 single-choice decisions. These made up our test (see Appendix B).

As might have been anticipated, the testing instrument did not have items evenly distributed among choices involving semantic meaning, structure of word groups and sentences, and extended discourse; the largest part of the test was devoted to items on structure, with the remainder devoted to semantic meaning and extended discourse. Further, test items would be shown to be invalid because of vague wording and the like. However we proceeded with the testing, which called for giving the test first to a criterion group of 21 people who are acknowledged experienced users of language in their vocations or in their daily work lives. The group ranged from students to professional writers, from secretaries to writing teachers. We specifically avoided a predominance of college teachers of English because we felt their responses might reflect only one dialect of our language, written standard English, and, further, because many had had training which would allow them to see deep structure as an intellectual concept, thus clouding the intuitive response the test was designed to ellicit. The responses of the criterion group became the basis for judging "right" and "wrong" answers in the main group of subjects, 158 students enrolled in our junior-level writing classes. Because of lack of agreement by members of the criterion group our stock of items was to be further reduced at the time of analysis. The standard applied was that 16 responses out of the 21 of the group should be in agreement on any specific item. Disagreement appeared to come from ambiguities in the instrument, some confusion over deep structure (a concept unfamiliar to most of the group), and variability of responses with items that attempted to reach overly fine discriminations. Requiring, then, the agreement of 16 out of 21 for each item, the test was finally reduced from 73 to 52 items for purposes of analysis. Meanwhile, the 73 item test was given to the

experimental group, composed of junior-senior students in Writing 323.

The six writing classes were taught, three each, by Williams and Nattinger, with parallel texts, parallel readings, parallel paper assignments and exercises, and, in so far as was possible, the same styles of classroom instruction. The chief purpose of instruction, in this case, was clarification of the idea of deep structure and establishment of some common store of knowledge about language in general. The last was most important because we had no way of knowing what the background of students would be. As it turned out, we had little difficulty in explaining the notion of deep structure, apparently, and we proceeded in each class to administer a pre- and a post-test. The pre-test was given the middle of the second week of instruction, the post-test was given at final examination time. In both testings, students were advised that no grade was attached to results, and, except for the fact of there being two testings, were informed what the project was about. Actual beginning enrollment in the courses totalled 246, but dropping of a class, failure to take both the pre- and post-test, and other causes resulted in a decreased subject population, 158. While this reduction was more than we had hoped would occur, the limited scope of the project made its impact less serious than it would be in a fully developed study.

RESULTS

The overall results of analysis of testing point toward six areas of measurable responsiveness, response to (1) sociolinguistic context, (2) semantic context, (3) deep structure of phrases, (4) deep structure of sentences, (5) primary semantic factors of extended discourse, and (6) organization of discourse.

Sociolinguistic context refers to the social setting in which language occurs, and involves the nature of a speaker and similar factors; it is not to be equated with "propriety," since many text items ranged over different levels and styles of expression and involved formal as well as informal social situations. Semantic context means, more narrowly, the connections with and influences on a given word or expression by surrounding words and meanings. It differs from "semantic meaning" (in the sense of "dictionary definition") by emphasis on the modifications of a word or expression brought about by the linguistic surrounding in which it occurs.

The two areas of deep structure are merely different ranges of the same phenomenon, and their meanings as we use them refer to the fundamental concept of deep structure as distinct from surface structure. Items were designed to test a subject's response to complex connections underlying phrase and sentence formulation, connections which go deeper than "agreement between subject and verb" or other superficial arrangements we observe as users of the language. The distinction, then, between the two areas is mainly one of the size of the linguistic unit, i.e., phrase or word-group, and sentence.

Primary semantic factors of extended discourse (listed on our factoring sheets as "Primary Factors of Extended Discussion," the last word being a misreading for "discourse") refers to the basic semantic component of discourse, discourse being any unit of verbal communication larger than the sentence. As with semantic context primary semantic factors involves more than the "dictionary definition" of a word; it includes the basic meaning of a word and the modifications of that meaning through the influence of surrounding linguistic elements. This individual and contextual meaning becomes, in extended discourse, further modified by structure of sentences and of paragraphs, but we treat it as a distinguishable element

for purposes of analysis. The same may be said of organization of discourse, which is, absolutely speaking, tied to the meanings of words, word groups, sentence form, and devices such as repetition, but which we treat as a separate factor. Organization of discourse includes such things as comparison-contrast, spacial arrangement, temporal arrangement, analogy, induction, deduction, cause and effect, and syllogistic form (See Appendix A).

These six areas are those which, out of a number we built into the test, survived analysis by experts and testing with a criterion and experimental group, and which emerged from analysis of results. Of the 73 items of the testing instrument (reduced by some disagreement among the criterion group) 52 items qualified for analysis and yielded the following distribution.

Nature of item	Number of items
Sociolinguistic Context	11
Semantic Context	7
Deep Structure, word-group	9
Deep Structure, sentence	12
Primary Semantic Factors of Extended Discourse	8
Organization of Extended Discourse	8
Number of items	52

Because three test items involved substantially more than one linguistic factor, the total number of items fed into computor analysis was 55. These six areas of apparently valid linguistic choice became, in effect, criteria of responsiveness the project was designed to search out.

However, the question of their reliability and validity remained, and further analysis was conducted to see to what extent items on the test registered differences of response and what these differences might reveal. In order to study reliability and validity, six scales were constructed, one each for the factors mentioned above. Since the attributes in question are dependent variables which presumably are amenable to change over time, test-retest reliability is inappropriate. Consequently pre-test and post-test data were analyzed separately, providing a form of replication. Odd-even split-half correlations were computed for each scale and corrected for reduction in length by the Spearman-Brown Prophecy Formula. As a further check on attenuation produced

by insufficient number of items, Scales A and B were combined (odd items from A plus even items from B versus even items from A plus odd items from B), Scales C and D were combined (same as A and B), and Scales E and F were combined (same as A and B).

Factor A, Sociolinguistic Context, consists of eleven items which have score weights ranging from 16 to 35 with a median of 28. Reliability estimates are a little low (.55 and .52) but could well be raised by further development.

Factor B, Semantic Context, consists of seven items having score weights ranging from 12 to 32 with a median of 28. Reliability estimates appear to be quite low (.14 and .21); this scale does not look very promising. Combining A and B yielded somewhat higher reliability coefficients (.67 and .61) but the increase is not as great as would be obtained by adding seven items to Scale A.

Factor C, Deep Structure (word-group), consists of nine items having score weights ranging from 29 to 42 with a median of 37. Reliability coefficients are about the same as for Factor A (.56 and .53).

Factor D, Deep Structure (sentence), consists of twelve items having score weights ranging from 27 to 43 with a median of 38. Reliability estimates are at the same level as for Factors A and C (.56 and .55). Combining Scales C and D has relatively little effect on reliability (.63 and .63).

Factor E, Primary Semantic Factors of Extended Discourse, consists of eight items having scoring weights ranging from 21 to 46 with a median of 30. Reliability is slightly lower than for Factors A, C and D (.50 and .47).

Factor F, Organization of Extended Discourse, consists of eight items having score weights ranging from 12 to 45; the item weights are 12, 40, 42, 45, 45, 45, 45, 45. Reliability coefficients are quite low (.05 and..31). Combining E and F produces reliability coefficients of .32 and .54.

Reviewing the reliability findings (Table 1) it appears that there is no advantage to combining scales (A and B, C and D, E and F). It also appears that Scales A, C, D and E are promising whereas Scales B and F seem less worthwhile.

Performance on each of the six factor scales was analyzed separately by means of classes nested in Instructors ANOVAR with two repeated measures (pre-post).

Since this study represents rather crude attempts to develop measures of some constructs derived from a particular theoretical approach to language, refined validation studies are rather premature. But it was considered useful to carry out some studies of this sort as a way of getting some hints as to the probable worth of the measures which had been developed. As displayed in Tables 3 through 8, it was found that there were student-gains in performance on Factors A, B, C and D by the end of the writing course and no pre-post change in performance on Scales E and F. Though these changes are as predicted, they are quite small in magnitude (Table 2). It is quite possible that this is partly a reflection of the insensitivity of the instrument indicated by the less than satisfactory reliabilities found for the six factor scales.

Table 1

Reliability* of Factors and Testing Occasion

Factor	Pre-Test	Post-Test
A	.55	.52
B	.14	.21
A & B	.67	.61
C	.56	.53
D	.56	.55
C & D	.63	.63
E	.50	.47
F	.05	.31
E & F	.32	.54

*Odd-even product-moment correlation corrected with Spearman-Brown Prophecy Formula; all 158 S's participated on both testing occasions.

Table 2

Mean Factor Scores

	FACTOR					
	A	B	C	D	E	F
Instructor 1	214	132	250*	340*	191	236
Instructor 2	224	133	271*	366*	200	234
Class 1	204	131	255	356	196	238
Class 2	205	130	248	332	177	234
Class 3	223	126	246	344	188	241
Class 4	224	138	250	344	201	231
Class 5	224	135	283	374	200	228
Class 6	225	135	277	373	203	237
Class 7	234	134	262	360	203	220
Class 8	212	130	264	356	192	249
Pre-Test	215*	129*	254*	344*	197	237
Post-Test	222*	135*	265*	359*	192	233

		CLASS							
		1	2	3	4	5	6	8	8
Factor A	Pre	194	192	227	230	217	220	227	212
	Post	214	218	219	218	231	230	241	212
Factor B	Pre	132	126	120	131	133	133	123	140
	Post	130	134	133	146	138	137	145	120
Factor C	Pre	254	250	220	245	281	273	255	263
	Post	256	246	272	254	284	281	269	265
Factor D	Pre	342	328	316	329	367	371	357	357
	Post	369	337	340	359	381	375	363	349

Table 3

FACTOR A ANOVAR

Source of Variation	SS	df	MS	F ratio	Sig
Instructor	6880	1	6880	1.91	ns
Class (Instructor)	23928	7	3418	0.95	ns
Error (Between)	535678	149	3595		
Total (Between)	566486	157			
Pre-Post	4221	1	4221	4.26	.01
Instructor x Pre-Post	229	1	229	0.23	ns
Error (Within)	147590	149	991		
Total (Within)	172659	158			

Table 4

FACTOR B ANOVAR

Source of Variation	SS	df	MS	F ratio	Sig
Instructor	280	1	280	0.24	ns
Class (Instructor)	4237	7	605	0.52	ns
Error (Between)	172439	149	1157		
Total (Between)	176956	157			
Pre-Post	2898	1	2898	7.51	.01
Instructor x Pre-Post	957	1	957	2.48	ns
Class (Instructor) x Pre-Post	10877	7	1554	4.03	.01
Error (Within)	57570	149	386		
Total (Within)	72302	158			

Table 5

FACTOR C

ANOVAR

Source of Variation	SS	df	MS	F ratio	Sig
Instructor	35056	1	35056	6.23	.01
Class (Instructor)	11608	7	1658	0.29	ns
Error (Between)	838224	149	5626		
Total (Between)	884888	157			
Pre-Post	9934	1	9934	6.57	.01
Instructor x Pre-Post	1246	1	1246	0.82	ns
Class (Instructor) x Pre-Post	57998	7	8285	5.48	.01
Error (Within)	225300	149	1512		
Total (Within)	294478	158			

Table 6

FACTOR D ANOVAR

Source of Variation	SS	df	MS	F Ratio	Sig
Instructor	51160	1	51160	8.35	.01
Class (Instructor)	28692	7	4099	0.67	ns
Error (Between)	913146	149	6128		
Total (Between)	992998	157			
Pre-Post	17972	1	17972	11.51	.01
Instructor x Pre-Post	6684	1	6684	4.28	.05
Class (Instructor) x Pre-Post	61540	7	8791	5.63	.01
Error (Within)	232652	149	1561		
Total (Within)	318848	158			

Table 7

FACTOR E ANOVAR

Source of Variation	SS	df	MS	F ratio	Sig
Instructor	5842	1	5842	2.18	ns
Class (Instructor)	16890	7	2413	0.90	ns
Error (Between)*	400016	149	2685		
Total (Between)	422748	157			
Pre-Post	1969	1	1969	1.87	ns
Instructor x Pre-Post	885	1	885	0.84	ns
Class (Instructor) x Pre-Post	10245	7	1464	1.39	ns
Error (Within)*	157268	149	1055		
Total (Within)	170367	158			

* Winer (1962) recommends pooling nested factor sums of squares with appropriate error terms when none of the nested factors reaches significance at the .10 level. This is of course a matter of controversy and in this particular analysis would make no difference in outcome.

Table 8

FACTOR F ANOVAR

Source of Variation	SS	df	MS	F ratio	Sig
Instructor	124	1	124	0.03	ns
Class (Instructor)	19336	7	2762	0.70	ns
Error (Between)*	585660	149	3931		
Total (Between)	605120	157			
Pre-Post	1164	1	1164	0.50	ns
Instructor x Pre-Post	80	1	80	0.03	ns
Class (Instructor) x Pre-Post	9964	7	1423	0.61	ns
Error (Within)*	347554	149	2333		
Total (Within)					

* Winer (1962) recommends pooling nested factor sums of squares with appropriate error terms when none of the nested factors reaches significance at the .10 level. This is of course a matter of controversy and in this particular analysis would make no difference in the outcome.

CONCLUSION

Taken together, the results of the reliability and validity studies indicate that scales A,B,C,D, and E merit further development. Discarding scale F may be premature since it may have been inappropriately employed and may prove worthwhile if tested on freshmen and sophomores rather than on juniors. These scales, and the criteria they represent, relate to individual items in the testing instrument given in Appendix B in the following manner.

Scale and Criterion of Responsiveness	Test Items Used in Final Analysis
1. Factor A, Sociolinguistic Context	1a,1b,1c,2a,2b, 2c,3a,3b,5a,5b,5c.
2. Factor B, Semantic Context	3a,3c,5b,6a,7a, 13c.
3. Factor C, Deep Structure (word-group)	16a,16b,16c,17a, 17b,17c,18a, 18b, 18c.
4. Factor D, Deep Structure (sentence)	10a,10b,11b,12a, 12b,12c,13a,13b, 14a,14b,15a,15b.
5. Factor E, Primary Semantic Factors of Extended Discourse	19a,19b,20a,20b, 23a,24a,24b,24c.
6. Factor F, Organization of Extended Discourse	22a,24c,25a,27a, 27b,27c,28a,28c.

For those interested in pursuing these criteria of responsiveness to English, actual test items should be consulted.

All in all, the researchers felt that a crude but indicative beginning has been made to define and measure aspects of language response not ordinarily considered and perhaps not even measured by standard instruments. It is hoped the criteria of responsiveness we have located will be further refined in the development of more sophisticated language tests than those now available.

LITERATURE CITED

Baratz,J. and Shuy,R., eds. Teaching Black Children to Read, Center for Applied Linguistics, Wash. D.C., 1969.

Bernstein, Basil, "Social Class and Linguistic Development: A Theory of Social Learning," in Education, Economy, and Society, ed. A.H. Halsey, N.Y., 1961.

Burling, R., Man's Many Voices, ch.11, N.Y., 1970.

Carroll, John B., "Notes on the Measurement of Achievement in Foreign Languages," Cambridge, Mass., 1954 (Mimeo.).

_____, Research on Teaching Foreign Languages, Publ. of Lang. Lab. of Univ. of Mich., vol. 3, Ann Arbor, 1961.

Deutsch, Martin P., "The Disadvantaged Child and the Learning Process," in Education in Depressed Areas, ed. A. Harry Passow, N.Y., 1963.

Hunt, Kellog W., "An Instrument to Measure Syntactic Maturity," Dept. of English, Florida State Univ., 1968.

Labov, William, A Study of the Non-Standard English of Negro and Puerto Rican Speakers in New York City, Center for Applied Linguistics, Wash. D.C., 1969 [1969a].

_____, "A Study of Non-Standard English," Center for Applied Linguistics, Wash. D.C., 1969 [1969b].

Shuy, R., ed., Social Dialects and Language Learning, National Council of Teachers of English, Champaign, Ill., 1964.

Smith, Vernon H., "Measuring Teacher Judgement in the Evaluation of Written Composition," Research in the Teaching of English, 3 (2), 1969, 181-195.

Tikofsky, Ronald S. et al., "Perception of Grammaticalness," Univ. Mich. Center Research on Lang. and Lang. Behavior, Report BR-6-1784, Ann Arbor, 1968.

Wolfram, Walt, "Sociolinguistic Premises and the Nature of Non-standard Dialects," Center for Applied Linguistics, Wash. D.C., 1969.

Wood, Susan, "An Evaluation of Published English Tests," Wisconsin-English-Language Arts Curriculum Project, Wisc. Dept. of Public Instr., Bulletin No. 144, 1968.

APPENDIX A

ITEM # ____ GENERAL DESCRIPTION OF ITEM SUBJECT ____
PARTS # ____

Aims:

Major Factors:

Minor or Negligable Factors:

Other:

Parts of Item					Factors		
E	D	C	B	A			
					Sociolinguistic Context (Register)	SEMANTIC FACTORS	
					Semantic Context		
					Connotation-Denotation		
					Formal Signals Influencing Semantics		
					Other (Specify)		
					Formal Signals (prep's, conj's)	STRUCTURAL FACTORS	Surface Structure
					Word Groups		
					Complete-Incomplete Stmt.		Syntax of Sentence
					Position-Position Emphasis		
					Joinery of Word Groups, S - Verb		
					Sentence Type-Form		
					Other (Specify)		
					Underlying Grammat'l Relat'ship		Deep Structure
					Relationship Between Sent.	FACTORS of EXTENDED DISCOURSE	
					Primary Factors of Ext'd Discussion		
					Comparison-Contrast		Organization of Discourse
					Spacial Arrangement		
					Temporal Arrangement		
					Association		
					Analogy		
					Induction		
					Deduction		
					Cause-Effect		
					Syllogistic Form		
					Other (Specify)		
					Larger Forms of Discourse (Par.)		
					Rhetorical Form		

APPENDIX B

LANGUAGE PREFERENCE TEST

The following test is designed to find out how your responses to the English language compare with those of people who have acknowledged sensitivity to its nuances. The test goes on no official record, nor does its ultimate result influence your course grade.

The majority of items you are asked to respond to pertain to written English. Every item or parts of it asks you to choose between two alternatives; every item must be marked, and an item not marked will be counted as "wrong." The test is timed. Items not completed will have to be counted as "wrong," so the most advisable procedure is to mark down what you instinctively feel is "right" of a particular case, rather than marking down a choice after long deliberation.

For purposes of linguistic study you are asked to fill in the information requested below. This information, along with your name and score, will be kept with the test for a month; after that time your test will be assigned a number and your name and information about you destroyed. Test results can be found out by consultation with the instructor during that month.

Name _____

Writing 323, Sec. _____

Sex _____

Home address (family) _____

Other languages spoken at home _____

To what extent? _____

Previous formal instruction in English language structure _____

_____

Current year in college _____

Special notes, if any, on the above:

CONTEST AND WORD MEANINGS

The sentence below has a blank space to be filled by a suitable expression. For each of the following pairs of possibilities, select the one which would be most suitable in filling in the blank.

Measured by the standards of his as well as our time, Abraham Lincoln was _____ leader of men.

A._____
1. a real persuasive
2. a most persuasive

B._____
1. a most persuasive
2. a really persuasive

C._____
1. a real persuasive
2. a really persuasive

Choose from each of the following pairs of possibilities the most likely user of the expression, "that's a no-no."

1. A minister talking to a parishioner about sin.

A.____ 2. An attorney tal ing to his son about using tools.

1. An operator of a beauty parlor talking to a customer about hair dye.

B.____ 2. A clerk talking to a customer about payment of a charge account.

1. A mechanic talking to his boss about how not to use an impact tool.

C.____ 2. An insurance adjuster speaking to his client about what not to say about a recent accident.

Choose from each pair of alternatives given below the one which is most suitable in filling the blank:

A.____ Our Savior Jesus was deeply loved by _____.

1. those who were his followers.
2. those individuals who followed him.

B.____ Every individual _____ a relationship with a certain amount of caution, especially when the relationship was with someone outside the group.

1. became involved in
2. got mixed up in

C.____ (same passage as in B)

1. entered into
2. participated in

A novelist wants to capture the feeling, the look and sense of autumn leaves in his description of a street. It is important that his story make the reader <u>feel</u>, in as many ways as possible, the "autumness" of the scene. For each of the following pairs of possibilities, select the one which best conveys the effect he wants.

1. The wind blew autumn leaves across the empty street.

A.____ 2. The wind blew parched autumn leaves across the empty street.

In a general discussion among neighbors on a warm afternoon over soft drinks in the backyard, the following are spoken. For each of the two possibilities given choose the one most likely and suitable to the situation.

A.____
1. "You'd never think, as such, that there would be any problem."
2. "You'd never think that there would be any problem, as such."

B.____
1. "So for that reason alone, one can never be certain of his fellow man."
2. "Thus, you can't be certain of your average man."

C.____
1. "It isn't a good field to play touch football on."
2. "It isn't a good field on which to play touch football."

30

Select the sentence most appropriate to the author's intent in each pair.

Young Stephen is caught up in the excitement of learning. The author wants to focus on the boy's awareness of his own ignorance, but also on the boy's immature pride in what he does know.

1. It pained him that he did not thoroughly know what politics meant and did not know where the universe ended.

A.____

2. It pained him that he did not know what politics meant and did not know where the universe ended.

1. It pained him that he did not know what politics meant and where the universe ended.

B.____

2. It pained him that he did not know what politics meant and did not know where the universe ended.

1. It pained him that he did not thoroughly know what politics meant and where the universe ended.

C.____

2. It pained him that he did not quite know what politics meant and where the universe ended.

31

Select the sentence most appropriate to the author's intent in each pair.

The author wants a statement to reflect his moral indignation against war.

1. It is men torturing and killing their enemies.

A.____

2. It is man torturing and killing his enemy.

1. It is man torturing and killing his enemy, and it is also the enemy being tortured and killed.

B.____

2. It is man torturing and killing his enemy, but it is also the enemy being tortured and killed.

32

Select the sentence most appropraite to the author's intent in each pair.

Maxwell Smith realizes he will have to borrow some money. The author wants the emphasis here to be from Smith's point of view, of his being in a situation where borrowing might be necessary.

A.____
1. He had never thought what it would be to borrow.
2. Borrowing was something he had never thought of doing before.

B.____
1. He had never thought what borrowing would be to him.
2. He had never thought what borrowing would be for him.

33

WORD COLLOCATIONS

Many times unusual combinations of words force us to modify our preconceived ideas about the meanings of one or the other of those words, e.g. "steel butterfly." If you came across the combination steel butterfly you would more likely change your ideas about butterflies than about steel. And it is not really a matter of which word comes first. Note the combination eloquent rain: it is most likely "eloquent" which would change meaning; "rain" remains unchanged.

For each pair below choose which word is most likely to change meaning when used in the underlined expression.

transparent chimpanzee

1. transparent

A.____

2. chimpanzee

brave coward

1. brave

B.____

2. coward

paralytic river

1. paralytic

C.____

2. river

34

UNDERLYING GRAMMATICAL RELATIONSHIPS

Many times we feel that certain sentences are closely related even though they look very different:

Class was dismissed by John
John dismissed class.

We have this feeling because we respond to the same underlying relationships: someone (John) did something (dismissed) to something (class). Likewise, we understand that the sentences.

John is eager to please
John is anxious to please

share the same underlying relationships (John wants to do something. i.e., John is the agent of the action), while another, which superficially looks exactly the same:

John is easy to please

is quite different in the essential relationship it expresses. Here, John is the object of the action, the thing acted upon.

From each of the pairs of alternatives below choose the one which comes closest to the underlined sentence in its underlying relationships.

<u>The window was broken by the hail.</u>

1. The window was broken by the garden.

A.____

2. The window was broken by the workman.

<u>The window was broken by the hail.</u>

1. The window was broken by 12 o'clock.

B.____

2. The window was broken by the falling branch.

See UNDERLYING GRAMMATICAL RELATIONSHIPS

The baby was too young to read to.

1. The baby was too young to talk.

A.____

2. The baby was too young to tickle.

The baby was too young to read to.

1. The baby was too young to sit.

B.____

2. The baby was too young to hold.

The baby was too young to read to.

1. The baby was too young to understand.

C.____

2. The baby was too young to punish.

See UNDERLYING GRAMMATICAL RELATIONSHIPS

She cooked him a roast turkey.

1. They made him a nervous wreck.

A.____

2. They built him a new home.

She cooked him a roast turkey.

1. They elected him president.

B.____

2. They sold him wheat.

She cooked him a roast turkey.

1. She made him a good wife.

C.____

2. She knitted him a blue sweater.

See UNDERLYING GRAMMATICAL RELATIONSHIPS

This is a riddle to puzzle everybody.

1. This is a riddle to ask everybody.

A.____

2. This is a riddle to entertain everybody.

This is a riddle to puzzle everybody.

1. This is a riddle to read to everybody.

B.____

2. This is a riddle to interest everybody.

This is a riddle to puzzle everybody.

1. This is a riddle to amuse everybody.

C.____

2. This is a riddle to please everybody.

See UNDERLYING GRAMMATICAL RELATIONSHIPS

They had an hour to read.

1. She had an article to read.

A.____

2. She had a reason to read.

They had an hour to read.

1. She had a good place to read.

B.____

2. She had a bad report to read.

They had an hour to read.

1. She had a book to read.

C.____

2. She had a copy to read.

37.

39

See UNDERLYING GRAMMATICAL RELATIONSHIPS

The car drove smoothly.

1. The shirt ironed neatly.

A.____

2. The branch fell suddenly.

The car drove smoothly.

1. The machine handles fairly well.

B.____

2. The shirt irons fairly well.

The car drove smoothly.

1. A woodsman sawed the tree.

C.____

2. A rock broke the stick.

COMPOUNDING

Many compounded phrases are built from the same kinds of underlying relationships. For example, in the following compounds:

(1) notary stamp
(2) book cover
(3) record player
(4) finger ring

Numbers (2) and (4) share the same underlying relations. This can be shown (and tested for) most readily by rephrasing each of the compounds. In (2), it's a cover on a book, and in (4), a ring on a finger, while (1) is perhaps the stamp of a notary, and (3), a player of/for records. Compounds that can undergo the same kinds of rephrasings usually share the same underlying relationships.

Which of the following noun compounds are formed from the same kinds of internal grammatical relations as the underlined compound.

41

A.____ brick-mason
1. water pistol
2. blood-donor

B.____ brick-mason
1. car-thief
2. cave-man

C.____ brick-mason
1. limestone
2. locksmith

See COMPOUNDING

ground-water

1. kidneystone

A.____

2. brick-mason

ground-water

1. bug-spray

B.____

2. space-charge

ground-water

1. body-fluids

C.____

2. gravel-pit

42

See COMPOUNDING

Ash-tray

1. briefcase

A.____

2. coffeecream

Ash-tray

1. fieldartillery

B.____

2. pigpen

Ash-tray

1. pocketbook

C.____

2. icebox

43

EXTENDED DISCOURSE

Choose from the alternatives given the ones which best fit the blank left in the passage below (the sense varies from choice to choice, but that is not important).

> The whole idea of the project was community involvement. That's why people objected. no one listened to what <u>they</u> wanted. _____ _____.

A.____ 1. The high rise apartment will never take the place of a home where people live and raise their family.

2. High rise apartments will never take the place of homes where people live and raise their families.

B.____ 1. An individual wants to decide where and how he's going to live.

2. The individual wants to decide where and how he's going to live.

44

From the alternatives below choose the one which fills the blank so as to provide continuity of idea in the following passage.

> The subject of the meeting was police relations with the community. _____
> _____
> _____. It was becoming dangerous, they said, to be on the streets at night. Of course part of the problem was poor street lighting.

A.____
1. Many people from the community objected to arrest procedures; all agreed that the area should be policed more frequently.
2. All agreed that the area should be policed more frequently; many people from the community objected to arrest procedures.

B.____
1. While many people from the community objected to arrest procedures, all agreed that the area should be policed more frequently.
2. While all agreed that the area should be policed more frequently, many people from the community objected to arrest procedures.

45

The writer of the following passage wants to add to it a comparison which will make his subject matter clearer to his reader. Of the two alternatives given choose the one which is most in keeping with the passage and the thought it expresses.

> Not only do maples adapt to climatic conditions by shedding leaves, they respond to heat and cold by thickness and flow of their sap: hence the saying, "the sap is up" or "the sap is down." _____
> _____
> _____
> _____.

A.____ 1. The maple's life is like a human being's in many respects: it has "blood," "arteries," "circulation" and different metabolisms or rates of body function.

2. The maple's life is like an insect's in many respects: it responds to heat and sum by searching for food and by propagating, and to cold and snow by slowing down its use of energy.

94

The writer of the following passage wants to split it into two paragraphs. Where would this split most naturally occur? Choose from the pairs of alternatives given below that which most nearly serves as a paragraph break. (Marginal numbers for reference).

The real art of taxation is in relieving people of
A-1 their wealth in as agreeable a way as possible. This
skill requires an approach which is at once courteous
A-2 and threatening, with emphasis on courtesy. For though
taxes will not be fully paid without threat, citizen
compliance is the cheapest method of gaining revenue;
were taxation to rely upon threat alone, a good share
B-1 of revenue gained would have to be spent on law enforcement. Much public relations effort is required of the
Internal Revenue Service to assure citizen compliance,
much taxpayer education is carried on by it to assure
B-2 that laws are known. For the Internal Revenue Service
the aim of public relations is not so different from
C-1 that of an advertising agency. In fact, one could
say that the IRS is selling one of the most difficult
"products" in our economy, the willful giving up of
C-2 one's money. And it is selling successfully, to judge
from the large share of taxes paid without undue regret.

A.____ 1. --as possible. This skill--

2. --on courtesy. For though--

B.____ 1. --law enforcement. Much public relations--

2. --laws are known. For the Internal--

C.____ 1. --advertising agency. In fact--

2. --one's money. And it is selling--

45.

47

The writer wants to delete some repetitive words (or terms) in the passage below. Of those underlined which are most repetitive and most likely to come in for change? Choose among alternatives given.

> Building codes tend to stifle creativity. They are inherently conservative, a codified accumulation of builders' experience which leaves little room for innovation. For instance, what if an architect wants over fifty percent of the area of a ceiling less than seven feet in height? He cannot have it because it would violate code. This is but one instance of thwarted imagination blocked by sheer inertia rather than by structural impossibility or occupant safety. Certainly if it were a matter of safety, the rigid provisions of building codes would be justified. And if new materials were not introduced, the codes' conservatism would be warranted, but almost daily...

A.____ 1. codified
2. inertia

B.____ 1. one instance
2. matter of safety

C.____ 1. matter of safety
2. conservatism

48

From each of the pairs of alternatives given choose the one which least belongs to the passage below, in terms of general suitability and appropriateness of meaning.

Normally our warranty does not cover parts which wear out quickly. With large orders, however, our organization stretches a point. fan belts, cutter blades, clutches, all can be replaced without cost to you. And I think you can see the reasoning behind that. We value your business, and we hope that AMSCO equipment stays in your production picture. We pride ourselves on being a vital part of America's food processing industry, one of the underpinnings of our economy. You will find our prices competitive and our equipment reliable, we mean it when we way, "AMSCO wants to join you."

A-1
C-2
B-1
A-2
B-2,C-1

A._____

1. fan belts, cutter blades, clutches, all can be replaced without cost to you.
2. one of the underpinnings of our economy.,

B._____

1. We pride ourselves on being a vital part of America's food processing industry, one of the underpinnings of our economy.
2. You will find our prices competitive and our equipment reliable, we mean it when we say, "AMSCO wants to join you."

C._____

1. You will find our prices competitive and our equipment reliable...
2. And I think you can see the reasoning behind that.

49

Read the passage below and then choose among the pairs of alternatives the ones which best describe what the passage emphasizes.

> In front lay the valley. And across it, the mountains stood defiantly, locking the land in stifling heat. To the left, from what appeared to be a volcanic fissure, a dried creek bed began to cut the valley floor, winding through the center of the land then back towards the mountains, finally, slowly, disappearing in the distance.

A.____ 1. Accumulation of descriptive details

2. Suspense (as in a detective story)

B.____ 1. Physical details described to give us a sense of the geograpny of the valley.

2. Revelation of a character's impressions of the valley.

50

From the alternatives given below choose one of each pair which best describes how the given passage is organized.

> The attempt to block the parking structure was well supported but too late. Even if the city council had wanted to stop building of it, applications had been submitted and approved in accordance with zoning regulations for that area. The proper time for complaint, citizens found, would have been at earlier hearings on zoning changes. Now the only thing that could be done was restriction of building height, and even that restriction could be cumvented.

A.____ 1. Narrative account.
2. Time sequence.

B.____ 1. Time sequence.
2. Cause and effect.

C.____ 1. Normal time sequence altered to emphasize cause and effect.
2. Cause and effect altered to fit time sequence.

51

From the pairs of alternatives given, choose the one which best describes the way the passage below is put together.

> Yellow, of course, is the color of wealth, as with gold. In fact, a certain shade of yellow was reserved for royalty in imperial China, but in this case the basic idea was sun worship. Just as the western church took over ancient pagan rites for its days of celebration, so the ruling class of China took advantage of worship of the sun. For us there is the same worship in our representations of the sun and sunshine, which are, ot this day, invariably yellow.

A.____ 1. Association of idea with idea.
2. Comparison and contrast.

B.____ 1. Cause and effect.
2. Comparison and contrast.

C.____ 1. Association of idea with idea.
2. Differentiation between things that appear alike.

50. 52

From each of the alternatives below choose the one of each pair which is most logical as statement. (Ultimate truth is not the question.)

1. All Republicans believe in free enterprise versus government control of the economy. Some Democrats, too, prefer that government be kept out of the country's economic affairs. That is why some Democrats are really Republicans.

A.____

2. All Republicans believe in free enterprise versus government control of the economy. Some Democrats, too, prefer that government be kept out of the country's economic affairs. That is why some Democrats are close to being Republicans.

1. This dude black at heart, man. I mean, he ain't white. Just about everybody either black or white when it come to a feeling bout race.

B.____

2. This dude black at heart, man. I mean, he ain't white. And everybody either black or white when it come to a feeling bout race.

51. 53

1. I'm really turning out work on this job, and if you work hard your pay will be better than last time, which is why I'll bet a better paycheck than I got then.

C.____

2. I'm really turning out work on this job, and if you work hard your pay will be better than last time, which is why I expect a better paycheck than I got then.